

Connection between MP and DPP for Stochastic Recursive Optimal Control Problems: Viscosity Solution Framework in Local Case

Tianyang Nie, Jingtao Shi, *Member, IEEE*, Zhen Wu

Abstract—This paper deals with a nonsmooth version of the connection between the maximum principle and dynamic programming principle, for the stochastic recursive control problem when the control domain is convex. By employing the notions of sub- and super-jets, the set inclusions are derived among the value function and the adjoint processes. The general case for non-convex control domain is open.

I. INTRODUCTION

There are usually two ways to study optimal control problems: Pontryagin's maximum principle (MP) and Bellman's dynamic programming principle (DPP), involving an adjoint variable ψ and the value function V , respectively. The classical result by Fleming and Rishel [7] on the connection between the MP and DPP is known as $\psi(t) = -V_x(t, \bar{x}(t))$, where $\bar{x}(\cdot)$ is the optimal state. Since the value function V is not always smooth, some non-smooth versions of the classical result are researched by non-smooth analysis and generalized derivatives. Within the framework of viscosity solution, Zhou [20] showed that

$$D_x^{1,-}V(t, \bar{x}(t)) \subset \{-\psi(t)\} \subset D_x^{1,+}V(t, \bar{x}(t)), \quad (1)$$

where $D_x^{1,-}V(t, \bar{x}(t))$ and $D_x^{1,+}V(t, \bar{x}(t))$ denote the first-order sub- and super-jets of V at $(t, \bar{x}(t))$, respectively.

For stochastic optimal control problems, the classical result on the connection between the MP and DPP is proved by Bensoussan [1], which is known as $p(t) = -V_x(t, \bar{x}(t))$, $q(t) = -V_{xx}(t, \bar{x}(t))\sigma(t, \bar{x}(t), \bar{u}(t))$ involving an adjoint process pair (p, q) , where $\bar{u}(\cdot)$ is the optimal control and σ is the diffusion coefficient. Within the framework of viscosity solution, Yong and Zhou [18] showed that

$$\begin{aligned} \{-p(t)\} \times [-P(t), \infty) &\subset D_x^{2,+}V(t, \bar{x}(t)), \\ D_x^{2,-}V(t, \bar{x}(t)) &\subset \{-p(t)\} \times (-\infty, -P(t)], \end{aligned} \quad (2)$$

where $D_x^{2,-}V(t, \bar{x}(t))$ and $D_x^{2,+}V(t, \bar{x}(t))$ denote the second-order sub- and super-jets of V at $(t, \bar{x}(t))$, and p, P are the first- and second-order adjoint processes, respectively.

In this paper, we consider one kind of stochastic recursive optimal control problem, where the cost functional is described by the solution to a *backward stochastic differential*

equation (BSDE) of the following form

$$\begin{cases} -dy(t) = f(t, y(t), z(t))dt - z(t)dW(t), & t \in [0, T], \\ y(T) = \xi, \end{cases}$$

where the terminal condition (rather than the initial condition) ξ is given in advance. Linear BSDE was introduced by Bismut [2], to represent the adjoint equation when applying the MP to solve stochastic optimal control problems. The nonlinear BSDE was introduced by Pardoux and Peng [9]. Independently, Duffie and Epstein [4] introduced BSDE from economic background, and they presented a stochastic differential formulation of recursive utility which is an extension of the standard additive utility with the instantaneous utility depending not only on the instantaneous consumption rate but also on the future utility. Stochastic recursive optimal control problems have found important applications in mathematical economics, mathematical finance and engineering (see El Karoui, Peng and Quenez [5], [6], Wang and Wu [15], Cvitanic and Zhang [3] and the references therein).

For stochastic recursive optimal control problems, Peng [11] first obtained a local maximum principle when the control domain is convex. And Xu [17] studied the non-convex control domain case, but with the assumption that the diffusion coefficient does not depend on the control variable. Wu [16] established a general maximum principle by Ekeland variational principle, where the control domain is non-convex and the diffusion coefficient contains the control variable. Peng [10] (also see Peng [12]) first obtained the generalized dynamic programming principle and introduced a generalized *Hamilton-Jacobi-Bellman* (HJB) equation which is a second-order parabolic *partial differential equation* (PDE). The value function is proved to be the viscosity solution to the generalized HJB equation.

The connection between MP and DPP for stochastic recursive optimal control problems was first studied by Shi [13] (see also Shi and Yu [14]) in its local form, when the control domain is convex and the value function is assumed to be smooth enough. The main result is

$$\begin{cases} p(t) = V_x(t, \bar{x}(t))^\top q(t), \\ k(t) = [V_{xx}(t, \bar{x}(t))\sigma(t, \bar{x}(t), \bar{u}(t)) + V_x(t, \bar{x}(t))f_z(t, \bar{x}(t), \\ -V(t, \bar{x}(t)), -V_x(t, \bar{x}(t))\sigma(t, \bar{x}(t), \bar{u}(t)), \bar{u}(t))]q(t), \end{cases} \quad (3)$$

involving an adjoint process triple (p, q, k) , where f is the generator of the controlled BSDE which is coupled with the controlled SDE. Applications to the recursive utility portfolio optimization problem in the financial market are discussed.

However, this classical result is highly unsatisfactory because the smoothness assumption on the value function V

This work was supported by the National Natural Science Foundations of China (11301011, 11201264, 11571205, 61573217), the 111 project (B12023), the Chang Jiang Scholar Program of Chinese Education Ministry, the Natural Science Foundations of Shandong Province of China (JQ201401, ZR2015JL003) and the Fundamental Research Fund of Shandong University (2015HW023).

Corresponding author: Jingtao Shi. T. Y. Nie, J. T. Shi and Z. Wu are all with School of Mathematics, Shandong University, Jinan 250100, P. R. China. nientianyang@sdu.edu.cn, shijingtiao@sdu.edu.cn, wuzhen@sdu.edu.cn

is illusory and it is not true even in the very simple case: see Example 3.1 of this paper. In the current work, we extend the above classical result by getting rid of the illusory assumption that the value function is differentiable. Our main contribution is to show the connection between the adjoint processes $p(\cdot), q(\cdot)$ in the maximum principle and the first-order sub- and super-jets $D_x^{1,-}V(t, \bar{x}(t)), D_x^{1,+}V(t, \bar{x}(t))$.

The rest of this paper is organized as follows. In Section 2, we state our problem and give some preliminary results about the MP and the DPP. Section 3 exhibits the main result of this paper, namely, the connection between the value function and the adjoint processes within the framework of viscosity solution. Finally, in Section 4 we give the concluding remarks.

II. PROBLEM STATEMENT AND PRELIMINARIES

Let $T > 0$ be finite and $\mathbf{U} \subset \mathbf{R}^k$ be nonempty and convex. Given $t \in [0, T]$, we denote $\mathcal{U}^w[t, T]$ the set of all 5-tuples $(\Omega, \mathcal{F}, \mathbf{P}, W(\cdot); u(\cdot))$ satisfying the following:

- (i) $(\Omega, \mathcal{F}, \mathbf{P})$ is a complete probability space;
- (ii) $\{W(s)\}_{s \geq t}$ is a d -dimensional standard Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{P})$ over $[t, T]$ (with $W(t) = 0$ almost surely), and $\mathcal{F}_s^t = \sigma\{W(r); t \leq r \leq s\}$ augmented by all the \mathbf{P} -null sets in \mathcal{F} ;
- (iii) $u : [t, T] \times \Omega \rightarrow \mathbf{U}$ is an $\{\mathcal{F}_s^t\}_{s \geq t}$ -adapted process on $(\Omega, \mathcal{F}, \mathbf{P})$.

We write $(\Omega, \mathcal{F}, \mathbf{P}, W(\cdot); u(\cdot)) \in \mathcal{U}^w[t, T]$, but occasionally we will write only $u(\cdot) \in \mathcal{U}^w[t, T]$ if no ambiguity exists. For any $(t, x) \in [0, T] \times \mathbf{R}^n$, consider the state $X^{t,x;u}(\cdot) \in \mathbf{R}^n$ given by the following controlled SDE:

$$\begin{cases} dX^{t,x;u}(s) = b(s, X^{t,x;u}(s), u(s))ds \\ \quad + \sigma(s, X^{t,x;u}(s), u(s))dW(s), \quad s \in [t, T], \\ X^{t,x;u}(t) = x. \end{cases} \quad (4)$$

Here $b : [0, T] \times \mathbf{R}^n \times \mathbf{U} \rightarrow \mathbf{R}^n, \sigma : [0, T] \times \mathbf{R}^n \times \mathbf{U} \rightarrow \mathbf{R}^{n \times d}$ are given functions. We assume that

(H1) b, σ are uniformly continuous in (s, x, u) , and there exists a constant $C > 0$ such that for all $s \in [0, T], x, \hat{x} \in \mathbf{R}^n, u \in \mathbf{U}$,

$$\begin{cases} |b(s, x, u) - b(s, \hat{x}, u)| + |\sigma(t, x, u) - \sigma(s, \hat{x}, u)| \leq C|x - \hat{x}|, \\ |b(s, x, u)| + |\sigma(s, x, u)| \leq C(1 + |x|). \end{cases}$$

For any $u(\cdot) \in \mathcal{U}^w[t, T]$, under **(H1)**, SDE (4) has a unique solution $X^{t,x;u}(\cdot)$ by the classical SDE theory (see [8], [18]). We refer to such $u(\cdot) \in \mathcal{U}^w[t, T]$ as an admissible control and $(X^{t,x;u}(\cdot), u(\cdot))$ as an admissible pair.

Next, we introduce the following controlled BSDE coupled with (4):

$$\begin{cases} -dY^{t,x;u}(s) = f(s, X^{t,x;u}(s), Y^{t,x;u}(s), Z^{t,x;u}(s), u(s))ds \\ \quad - Z^{t,x;u}(s)dW(s), \quad s \in [t, T], \\ Y^{t,x;u}(T) = \phi(X^{t,x;u}(T)). \end{cases} \quad (5)$$

Here $f : [0, T] \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^d \times \mathbf{U} \rightarrow \mathbf{R}, \phi : \mathbf{R}^n \rightarrow \mathbf{R}$ are given functions. We assume that

(H2) f, ϕ are uniformly continuous in (s, x, y, z, u) and there exists a constant $C > 0$ such that for all $s \in$

$$[0, T], x, \hat{x} \in \mathbf{R}^n, y, \hat{y} \in \mathbf{R}, z, \hat{z} \in \mathbf{R}^d, u \in \mathbf{U},$$

$$\begin{cases} |f(s, x, y, z, u) - f(s, \hat{x}, \hat{y}, \hat{z}, u)| \\ \leq C(|x - \hat{x}| + |y - \hat{y}| + |z - \hat{z}|), \\ |f(s, x, 0, 0, u)| + |\phi(x)| \leq C(1 + |x|), \\ |\phi(x) - \phi(\hat{x})| \leq C|x - \hat{x}|. \end{cases}$$

Then for any $u(\cdot) \in \mathcal{U}^w[t, T]$ and the given unique solution $X^{t,x;u}(\cdot)$ to (4), under **(H2)**, BSDE (5) admits a unique solution $(Y^{t,x;u}(\cdot), Z^{t,x;u}(\cdot))$ by the classical BSDE theory (see Pardoux and Peng [9] or Peng [12]).

Given $u(\cdot) \in \mathcal{U}^w[t, T]$, we introduce the cost functional

$$J(t, x; u(\cdot)) := -Y^{t,x;u}(s)|_{s=t}, \quad (t, x) \in [0, T] \times \mathbf{R}^n. \quad (6)$$

Our recursive stochastic optimal control problem is the following.

Problem (RSOCP). For given $(t, x) \in [0, T] \times \mathbf{R}^n$, to minimize (6) subject to (4)~(5) over $\mathcal{U}^w[t, T]$.

We define the value function

$$\begin{cases} V(t, x) := \inf_{u(\cdot) \in \mathcal{U}^w[t, T]} J(t, x; u(\cdot)), \quad (t, x) \in [0, T] \times \mathbf{R}^n, \\ V(T, x) = -\phi(x), \quad x \in \mathbf{R}^n. \end{cases} \quad (7)$$

Any $\bar{u}(\cdot) \in \mathcal{U}^w[t, T]$ that achieves the above infimum is called an optimal control, and the corresponding solution triple $(\bar{X}^{t,x;\bar{u}}(\cdot), \bar{Y}^{t,x;\bar{u}}(\cdot), \bar{Z}^{t,x;\bar{u}}(\cdot))$ is called an optimal state. We refer to $(\bar{X}^{t,x;\bar{u}}(\cdot), \bar{Y}^{t,x;\bar{u}}(\cdot), \bar{Z}^{t,x;\bar{u}}(\cdot), \bar{u}(\cdot))$ as an optimal quadruple.

Remark 2.1 Because b, σ, f, g are all deterministic functions, then from Proposition 5.1 of Peng [12], we know that under **(H1)**, **(H2)**, the above value function is a deterministic function. Thus our definition (7) is meaningful.

We introduce the following generalized HJB equation:

$$\begin{cases} -v_t(t, x) + \sup_{u \in \mathbf{U}} G(t, x, -v(t, x), -v_x(t, x), \\ \quad -v_{xx}(t, x), u) = 0, \quad (t, x) \in [0, T] \times \mathbf{R}^n, \\ v(T, x) = -\phi(x), \quad \forall x \in \mathbf{R}^n, \end{cases} \quad (8)$$

where the generalized Hamiltonian function $G : [0, T] \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{S}^n \times \mathbf{U} \rightarrow \mathbf{R}$ is defined as

$$\begin{aligned} G(t, x, r, p, A, u) &:= \frac{1}{2} \text{tr}\{\sigma(t, x, u)^\top A \sigma(t, x, u)\} \\ &\quad + \langle p, b(t, x, u) \rangle + f(t, x, r, \sigma(t, x, u)^\top p, u). \end{aligned} \quad (9)$$

The following result belongs to Peng [12].

Proposition 2.1 *Let **(H1)**, **(H2)** hold. Then for any $t \in [0, T]$ and $x, x' \in \mathbf{R}^n$, we have*

$$\begin{aligned} (i) \quad &|V(t, x) - V(t, x')| \leq C|x - x'|, \\ (ii) \quad &|V(t, x)| \leq C(1 + |x|). \end{aligned} \quad (10)$$

We introduce the definition of the viscosity solution for HJB equation (8).

Definition 2.1 (i) A function $v \in C([0, T] \times \mathbf{R}^n)$ is called a viscosity subsolution to (8) if

$$v(T, x) \leq -\phi(x), \quad \forall x \in \mathbf{R}^n,$$

and for any $\varphi \in C^{1,2}([0, T] \times \mathbf{R}^n)$, whenever $v - \varphi$ attains a local maximum at $(t, x) \in [0, T] \times \mathbf{R}^n$, we have

$$-\varphi_t(t, x) + \sup_{u \in \mathbf{U}} G(t, x, -v(t, x), -\varphi_x(t, x), -\varphi_{xx}(t, x), u) \leq 0.$$

(ii) A function $v \in C([0, T] \times \mathbb{R}^n)$ is called a viscosity supersolution to (8) if

$$v(T, x) \geq -\phi(x), \quad \forall x \in \mathbb{R}^n,$$

and for any $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^n)$, whenever $v - \varphi$ attains a local minimum at $(t, x) \in [0, T] \times \mathbb{R}^n$, we have

$$-\varphi_t(t, x) + \sup_{u \in \mathbf{U}} G(t, x, -v(t, x), -\varphi_x(t, x), -\varphi_{xx}(t, x), u) \geq 0.$$

(iii) A function $v \in C([0, T] \times \mathbb{R}^n)$ is called a viscosity solution to (8) if it is both a viscosity subsolution and viscosity supersolution to (8).

The following result also belongs to Peng [12].

Proposition 2.2 Let (H1), (H2) hold. Then $V(\cdot, \cdot)$ defined by (7) is the unique viscosity solution to (8).

To conveniently state the maximum principle, we regard the above (4), (5) as a controlled forward-backward stochastic differential equation (FBSDE):

$$\begin{cases} dX^{t,x;u}(s) = b(s, X^{t,x;u}(s), u(s))ds \\ \quad + \sigma(s, X^{t,x;u}(s), u(s))dW(s), \\ -dY^{t,x;u}(s) = f(s, X^{t,x;u}(s), Y^{t,x;u}(s), Z^{t,x;u}(s), u(s))ds \\ \quad - Z^{t,x;u}(s)dW(s), \quad s \in [t, T], \\ X^{t,x;u}(t) = x, \quad Y^{t,x;u}(T) = \phi(X^{t,x;u}(T)). \end{cases} \quad (11)$$

We need the following assumption.

(H3) b, σ, ϕ, f are continuously differentiable in (x, y, z) and the partial derivatives are uniformly bounded.

Let $(\bar{X}^{t,x;\bar{u}}(\cdot), \bar{Y}^{t,x;\bar{u}}(\cdot), \bar{Z}^{t,x;\bar{u}}(\cdot), \bar{u}(\cdot))$ be an optimal quadruple. For all $s \in [0, T]$, we denote

$$\begin{aligned} \bar{b}(s) &:= b(s, \bar{X}^{t,x;\bar{u}}(s), \bar{u}(s)), \quad \bar{\sigma}(s) := \sigma(s, \bar{X}^{t,x;\bar{u}}(s), \bar{u}(s)), \\ \bar{f}(s) &:= f(s, \bar{X}^{t,x;\bar{u}}(s), \bar{Y}^{t,x;\bar{u}}(s), \bar{Z}^{t,x;\bar{u}}(s), \bar{u}(s)), \end{aligned}$$

and similar notations are used for all their derivatives.

We introduce the adjoint equation:

$$\begin{cases} -dp(s) = [\bar{b}_x(s)^\top p(s) - \bar{f}_x(s)^\top q(s) + \bar{\sigma}_x(s)k(s)]ds \\ \quad - k(s)dW(s), \\ dq(s) = \bar{f}_y(s)^\top q(s)ds + \bar{f}_z(s)^\top q(s)dW(s), \quad s \in [t, T], \\ p(T) = -\phi_x(\bar{X}^{t,x;\bar{u}}(T))^\top q(T), \quad q(t) = 1, \end{cases} \quad (12)$$

and the Hamiltonian function $H : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \times \mathbf{U} \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{n \times d} \rightarrow \mathbb{R}$ is defined as

$$\begin{aligned} H(t, x, y, z, u, p, q, k) &:= \langle p, b(t, x, u) \rangle \\ &\quad - \langle q, f(t, x, y, z, u) \rangle + \text{tr}[\sigma(t, x, u)^\top k]. \end{aligned} \quad (13)$$

Under (H1), (H2), (H3), (12) admits a unique solution $(p(\cdot), q(\cdot), k(\cdot))$, which is called the adjoint process triple.

The following result comes from Peng [11].

Proposition 2.3 Let (H1), (H2), (H3) hold and $(t, x) \in [0, T] \times \mathbb{R}^n$ be fixed. Suppose that $\bar{u}(\cdot)$ is an optimal control for Problem (RSOCP), and $(\bar{X}^{t,x;\bar{u}}(\cdot), \bar{Y}^{t,x;\bar{u}}(\cdot), \bar{Z}^{t,x;\bar{u}}(\cdot))$ is the corresponding optimal state. Let $(p(\cdot), q(\cdot), k(\cdot))$ be the adjoint process triple. Then

$$\begin{aligned} \langle H_u(s, \bar{X}^{t,x;\bar{u}}(s), \bar{Y}^{t,x;\bar{u}}(s), \bar{Z}^{t,x;\bar{u}}(s), \bar{u}(s), \\ p(s), q(s), k(s)), u - \bar{u}(s) \rangle \geq 0, \quad \forall u \in \mathbf{U}, \end{aligned} \quad (14)$$

a.e. $s \in [t, T]$, \mathbf{P} -a.s.

Remark 2.2 Notice that Proposition 2.3 is proved by Peng [11] in its strong formulation. However, as pointed out in Yong and Zhou [18], since the DPP is involved, we need to deal with Problem (RSOCP) in its weak formulation. Since only necessary conditions

of optimality are considered here, an optimal quadruple (no matter whether in the strong or weak formulation) is given as a starting point, and all the results are valid for this given optimal quadruple on the probability space it attached to.

III. MAIN RESULT

We first introduce the notion of the first-order super- and sub-jets. For $v \in C([0, T] \times \mathbb{R}^n)$, and $(t, \hat{x}) \in [0, T] \times \mathbb{R}^n$, we define

$$\begin{cases} D_x^{1,+} v(t, \hat{x}) := \left\{ p \in \mathbb{R}^n \mid v(t, x) \leq v(t, \hat{x}) + \langle p, x - \hat{x} \rangle \right. \\ \quad \left. + o(|x - \hat{x}|), \text{ as } x \rightarrow \hat{x} \right\}, \\ D_x^{1,-} v(t, \hat{x}) := \left\{ p \in \mathbb{R}^n \mid v(t, x) \geq v(t, \hat{x}) + \langle p, x - \hat{x} \rangle \right. \\ \quad \left. + o(|x - \hat{x}|), \text{ as } x \rightarrow \hat{x} \right\}, \end{cases} \quad (15)$$

Theorem 3.1 Let (H1), (H2), (H3) hold and $(t, x) \in [0, T] \times \mathbb{R}^n$ be fixed. Suppose that $\bar{u}(\cdot)$ is an optimal control for Problem (RSOCP), and $(\bar{X}^{t,x;\bar{u}}(\cdot), \bar{Y}^{t,x;\bar{u}}(\cdot), \bar{Z}^{t,x;\bar{u}}(\cdot))$ is the corresponding optimal state. Let $(p(\cdot), q(\cdot), k(\cdot))$ be the adjoint process triple. Then

$$\begin{aligned} D_x^{1,-} V(s, \bar{X}^{t,x;\bar{u}}(s)) &\subset \{p(s)q^{-1}(s)\} \\ &\subset D_x^{1,+} V(s, \bar{X}^{t,x;\bar{u}}(s)), \quad \forall s \in [t, T], \mathbf{P}\text{-a.s.} \end{aligned} \quad (16)$$

where $V(\cdot, \cdot)$ is the value function defined by (7).

Proof. Fix an $s \in [t, T]$. For any $x^1 \in \mathbb{R}^n$, denote by $(X^{s,x^1;\bar{u}}(\cdot), Y^{s,x^1;\bar{u}}(\cdot), Z^{s,x^1;\bar{u}}(\cdot))$ the solution to the following FBSDE on $[s, T]$:

$$\begin{cases} X^{s,x^1;\bar{u}}(r) = x^1 + \int_s^r b(\alpha, X^{s,x^1;\bar{u}}(\alpha), \bar{u}(\alpha))d\alpha \\ \quad + \int_s^r \sigma(\alpha, X^{s,x^1;\bar{u}}(\alpha), \bar{u}(\alpha))dW(\alpha), \\ Y^{s,x^1;\bar{u}}(r) = \phi(X^{s,x^1;\bar{u}}(T)) + \int_r^T f(\alpha, X^{s,x^1;\bar{u}}(\alpha), \\ \quad Y^{s,x^1;\bar{u}}(\alpha), Z^{s,x^1;\bar{u}}(\alpha), \bar{u}(\alpha))d\alpha \\ \quad - \int_r^T Z^{s,x^1;\bar{u}}(\alpha)dW(\alpha), \quad r \in [s, T]. \end{cases} \quad (17)$$

It is clear that (17) can be regarded as an FBSDE on $(\Omega, \mathcal{F}, \{\mathcal{F}_r^t\}_{r \geq t}, \mathbf{P}(\cdot | \mathcal{F}_s^t)(\omega))$ for \mathbf{P} -a.s. ω , where $\mathbf{P}(\cdot | \mathcal{F}_s^t)(\omega)$ is the regular conditional probability given \mathcal{F}_s^t defined on (Ω, \mathcal{F}) .

For any $s \leq r \leq T$, set

$$\begin{aligned} \hat{X}(r) &:= X^{s,x^1;\bar{u}}(r) - \bar{X}^{t,x;\bar{u}}(r), \\ \hat{Y}(r) &:= Y^{s,x^1;\bar{u}}(r) - \bar{Y}^{t,x;\bar{u}}(r), \\ \hat{Z}(r) &:= Z^{s,x^1;\bar{u}}(r) - \bar{Z}^{t,x;\bar{u}}(r). \end{aligned}$$

Thus by a standard argument (see Theorem 6.3, Chapter 1, Yong and Zhou [18]), we have for any integer $k \geq 1$,

$$\mathbb{E} \left[\sup_{s \leq r \leq T} |\hat{X}(r)|^{2k} \middle| \mathcal{F}_s^t \right] \leq C |x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2k}, \quad \mathbf{P}\text{-a.s.} \quad (18)$$

Moreover, the following estimates holds by Peng [12],

$$\begin{aligned} \mathbb{E} \left[\sup_{s \leq r \leq T} |\hat{Y}(r)|^{2k} \middle| \mathcal{F}_s^t \right] &\leq C |x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2k}, \quad \mathbf{P}\text{-a.s.}, \\ \mathbb{E} \left[\left(\int_s^T |\hat{Z}(r)|^2 dr \right)^k \middle| \mathcal{F}_s^t \right] &\leq C |x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2k}, \quad \mathbf{P}\text{-a.s.} \end{aligned} \quad (19)$$

Now we write the equation for $\hat{X}(\cdot)$ as

$$\begin{cases} d\hat{X}(r) = \{\bar{b}_x(r)\hat{X}(r) + \varepsilon_1(r)\}dr \\ \quad + \sum_{j=1}^d \{\bar{\sigma}_x^j(r)\hat{X}(r) + \varepsilon_2^j(r)\}dW^j(r), r \in [s, T], \\ \hat{X}(s) = x^1 - \bar{X}^{t,x;\bar{u}}(s), \end{cases} \quad (20)$$

and the equation for $(\hat{Y}(\cdot), \hat{Z}(\cdot))$ as

$$\begin{cases} -d\hat{Y}(r) = \{\bar{f}_x(r)\hat{X}(r) + \bar{f}_y(r)\hat{Y}(r) + \bar{f}_z(r)\hat{Z}(r) \\ \quad + \varepsilon_3(r)\}dr - \hat{Z}(r)dW(r), r \in [s, T], \\ \hat{Y}(T) = \phi_x(\bar{X}^{t,x;\bar{u}}(T))\hat{X}(T) + \varepsilon_4(T), \end{cases} \quad (21)$$

respectively, where

$$\begin{cases} \varepsilon_1(r) := \int_0^1 [b_x(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{u}(r)) \\ \quad - \bar{b}_x(r)]\hat{X}(r)d\theta, \\ \varepsilon_2^j(r) := \int_0^1 [\sigma_x^j(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{u}(r)) \\ \quad - \bar{\sigma}_x^j(r)]\hat{X}(r)d\theta, j = 1, 2, \dots, d, \\ \varepsilon_3(r) := \int_0^1 [f_x(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{Y}^{t,x;\bar{u}}(r) + \theta\hat{Y}(r), \\ \quad \bar{Z}^{t,x;\bar{u}}(r) + \theta\hat{Z}(r), \bar{u}(r)) - \bar{f}_x(r)]\hat{X}(r)d\theta \\ \quad + \int_0^1 [f_y(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{Y}^{t,x;\bar{u}}(r) + \theta\hat{Y}(r), \\ \quad \bar{Z}^{t,x;\bar{u}}(r) + \theta\hat{Z}(r), \bar{u}(r)) - \bar{f}_y(r)]\hat{Y}(r)d\theta \\ \quad + \int_0^1 [f_z(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{Y}^{t,x;\bar{u}}(r) + \theta\hat{Y}(r), \\ \quad \bar{Z}^{t,x;\bar{u}}(r) + \theta\hat{Z}(r), \bar{u}(r)) - \bar{f}_z(r)]\hat{Z}(r)d\theta, \\ \varepsilon_4(T) := \int_0^1 [\phi_x(\bar{X}^{t,x;\bar{u}}(T) + \theta\hat{X}(T)) \\ \quad - \phi_x(\bar{X}^{t,x;\bar{u}}(T))]\hat{X}(T)d\theta. \end{cases}$$

As in pp. 258, Section 4, Chapter 5 of Yong and Zhou [18], for any $k \geq 1$, there exists a deterministic continuous and increasing function $\delta : [0, \infty) \rightarrow [0, \infty)$, independent of $x^1 \in \mathbf{R}^n$, with $\frac{\delta(r)}{r} \rightarrow 0$ as $r \rightarrow 0$, such that

$$\begin{cases} \mathbb{E}\left[\int_s^T |\varepsilon_1(r)|^{2k} dr | \mathcal{F}_s^t\right] \leq \delta(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2k}), \mathbf{P}\text{-a.s.}, \\ \mathbb{E}\left[\int_s^T |\varepsilon_2(r)|^{2k} dr | \mathcal{F}_s^t\right] \leq \delta(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2k}), \mathbf{P}\text{-a.s.}, \\ \mathbb{E}\left[|\varepsilon_4(T)|^{2k} | \mathcal{F}_s^t\right] \leq \delta(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2k}), \mathbf{P}\text{-a.s.} \end{cases} \quad (22)$$

Moreover, for some $0 < \alpha < 1$, we have

$$\mathbb{E}\left[\int_s^T |\varepsilon_3(r)|^{1+\alpha} dr | \mathcal{F}_s^t\right] \leq \delta(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{1+\alpha}), \mathbf{P}\text{-a.s.} \quad (23)$$

In fact, denote

$$\begin{cases} \Delta f_x(\theta) := f_x(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{Y}^{t,x;\bar{u}}(r) + \theta\hat{Y}(r), \\ \quad \bar{Z}^{t,x;\bar{u}}(r) + \theta\hat{Z}(r), \bar{u}(r)) - \bar{f}_x(r), \\ \Delta f_y(\theta) := f_y(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{Y}^{t,x;\bar{u}}(r) + \theta\hat{Y}(r), \\ \quad \bar{Z}^{t,x;\bar{u}}(r) + \theta\hat{Z}(r), \bar{u}(r)) - \bar{f}_y(r), \\ \Delta f_z(\theta) := f_z(r, \bar{X}^{t,x;\bar{u}}(r) + \theta\hat{X}(r), \bar{Y}^{t,x;\bar{u}}(r) + \theta\hat{Y}(r), \\ \quad \bar{Z}^{t,x;\bar{u}}(r) + \theta\hat{Z}(r), \bar{u}(r)) - \bar{f}_z(r). \end{cases}$$

Then

$$\begin{aligned} & \mathbb{E}\left[\int_s^T |\varepsilon_3(r)|^{1+\alpha} dr | \mathcal{F}_s^t\right] \\ & \leq 3\mathbb{E}\left[\int_s^T \left|\int_0^1 \Delta f_x(\theta)d\theta\right|^{1+\alpha} |\hat{X}(r)|^{1+\alpha} dr | \mathcal{F}_s^t\right] \\ & \quad + 3\mathbb{E}\left[\int_s^T \left|\int_0^1 \Delta f_y(\theta)d\theta\right|^{1+\alpha} |\hat{Y}(r)|^{1+\alpha} dr | \mathcal{F}_s^t\right] \\ & \quad + 3\mathbb{E}\left[\int_s^T \left|\int_0^1 \Delta f_z(\theta)d\theta\right|^{1+\alpha} |\hat{Z}(r)|^{1+\alpha} dr | \mathcal{F}_s^t\right] \\ & := 3I_1 + 3I_2 + 3I_3. \end{aligned}$$

We consider I_3 only. In fact, by Hölder's inequality, for $p = \frac{2}{1-\alpha}$, $q = \frac{2}{1+\alpha}$, we have

$$\begin{aligned} I_3 &= \mathbb{E}\left[\int_s^T \left|\int_0^1 \Delta f_z(\theta)d\theta\right|^{1+\alpha} |\hat{Z}(r)|^{1+\alpha} dr | \mathcal{F}_s^t\right] \\ & \leq \mathbb{E}\left[\left(\int_s^T \left|\int_0^1 \Delta f_z(\theta)d\theta\right|^{(1+\alpha)p} dr\right)^{\frac{1}{p}}\right. \\ & \quad \left.\left(\int_s^T |\hat{Z}(r)|^2 dr\right)^{\frac{1+\alpha}{2}} | \mathcal{F}_s^t\right] \\ & \leq \left\{\mathbb{E}\left[\left(\int_s^T \left|\int_0^1 \Delta f_z(\theta)d\theta\right|^{(1+\alpha)p} dr\right)^{\frac{2}{p}} | \mathcal{F}_s^t\right]\right\}^{\frac{1}{2}} \\ & \quad \left\{\mathbb{E}\left[\left(\int_s^T |\hat{Z}(r)|^2 dr\right)^{1+\alpha} | \mathcal{F}_s^t\right]\right\}^{\frac{1}{2}} \\ & := \Pi(\hat{X}(s))|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{1+\alpha}, \end{aligned}$$

since by the second inequality of (19), we have

$$\mathbb{E}\left[\left(\int_s^T |\hat{Z}(r)|^2 dr\right)^{1+\alpha} | \mathcal{F}_s^t\right] \leq C|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{2(1+\alpha)},$$

where

$$\Pi(\hat{X}(s)) := C\left\{\mathbb{E}\left[\left(\int_s^T \left|\int_0^1 \Delta f_z(\theta)d\theta\right|^{(1+\alpha)p} dr\right)^{\frac{2}{p}} | \mathcal{F}_s^t\right]\right\}^{\frac{1}{2}}.$$

Since from (H3) we have that $\Delta f_z(\cdot)$ is bounded and f_z is continuous, then from dominate convergence theorem, we have $\Pi(\hat{X}(s)) \rightarrow 0$, as $x^1 - \bar{X}^{t,x;\bar{u}}(s) \rightarrow 0$. That is, $I_3 \leq \delta(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|^{1+\alpha})$.

Similarly, by (18) and the first inequality of (19), we can obtain the same estimates for I_1, I_2 . Thus (23) holds.

Applying Itô's formula to $\langle \hat{X}(\cdot), p(\cdot) \rangle + \hat{Y}(\cdot)q(\cdot)$, noting (12), (20) and (21), we have

$$\begin{aligned} & \hat{Y}(s)q(s) = -\langle \hat{X}(s), p(s) \rangle + \mathbb{E}[\varepsilon_4(T)q(T) | \mathcal{F}_s^t] \\ & - \mathbb{E}\left[\int_s^T \langle \varepsilon_1(r), p(r) \rangle dr | \mathcal{F}_s^t\right] - \mathbb{E}\left[\int_s^T \langle \varepsilon_2(r), k(r) \rangle dr | \mathcal{F}_s^t\right] \\ & - \mathbb{E}\left[\int_s^T \varepsilon_3(r)q(r) dr | \mathcal{F}_s^t\right], \quad \mathbf{P}\text{-a.s.} \end{aligned} \quad (24)$$

Noting (22) and (23), since $\mathbb{E}\left[\sup_{s \leq r \leq T} |p(r)|^{2k} | \mathcal{F}_s^t\right] < \infty$, $\mathbb{E}\left[\sup_{s \leq r \leq T} |q(r)|^{2k} | \mathcal{F}_s^t\right] < \infty$, $\mathbb{E}\left[\int_s^T |k(r)|^2 dr | \mathcal{F}_s^t\right] < \infty$, it follows that

$$\begin{aligned} \mathbb{E}[\varepsilon_4(T)q(T) | \mathcal{F}_s^t] & \leq \left(\mathbb{E}[|\varepsilon_4(T)|^2 | \mathcal{F}_s^t]\right)^{\frac{1}{2}} \left(\mathbb{E}[|q(T)|^2 | \mathcal{F}_s^t]\right)^{\frac{1}{2}} \\ & \leq o(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|), \end{aligned}$$

$$\begin{aligned}
\mathbb{E} \left[\int_s^T \langle \varepsilon_1(r), p(r) \rangle dr \middle| \mathcal{F}_s^t \right] &\leq \mathbb{E} \left[\sup_{s \leq r \leq T} p(r) \int_s^T \varepsilon_1(r) dr \middle| \mathcal{F}_s^t \right] \\
&\leq \left(\mathbb{E} \left[\sup_{s \leq r \leq T} |p(r)|^2 \middle| \mathcal{F}_s^t \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\left(\int_s^T \varepsilon_1(r) dr \right)^2 \middle| \mathcal{F}_s^t \right] \right)^{\frac{1}{2}} \\
&\leq o(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|), \\
\mathbb{E} \left[\int_s^T \langle \varepsilon_3(r), k(r) \rangle dr \middle| \mathcal{F}_s^t \right] &\leq \left(\mathbb{E} \left[\left(\int_s^T k(r) dr \right)^2 \middle| \mathcal{F}_s^t \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\left(\int_s^T \varepsilon_3(r) dr \right)^2 \middle| \mathcal{F}_s^t \right] \right)^{\frac{1}{2}} \\
&\leq o(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E} \left[\int_s^T \varepsilon_3(r) q(r) dr \middle| \mathcal{F}_s^t \right] &\leq \mathbb{E} \left[\sup_{s \leq r \leq T} q(r) \int_s^T \varepsilon_3(r) dr \middle| \mathcal{F}_s^t \right] \\
&\leq \left(\mathbb{E} \left[\sup_{s \leq r \leq T} |q(r)|^q \middle| \mathcal{F}_s^t \right] \right)^{\frac{1}{q}} \left(\mathbb{E} \left[\left(\int_s^T \varepsilon_3(r) dr \right)^{1+\alpha} \middle| \mathcal{F}_s^t \right] \right)^{\frac{1}{1+\alpha}} \\
&\leq o(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|),
\end{aligned}$$

where $q = \frac{1+\alpha}{\alpha}$. Thus, we have

$$\hat{Y}(s)q(s) = -\langle \hat{X}(s), p(s) \rangle + o(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|), \quad \mathbf{P}\text{-a.s.} \quad (25)$$

Since $q(\cdot)$ is invertible, then

$$\hat{Y}(s) = -\langle \hat{X}(s), p(s)q^{-1}(s) \rangle + o(|x^1 - \bar{X}^{t,x;\bar{u}}(s)|), \quad \mathbf{P}\text{-a.s.} \quad (26)$$

Let us call a $x^1 \in \mathbf{R}^n$ rational if all its coordinates are rational numbers. Since the set of all rational $x^1 \in \mathbf{R}^n$ is countable, we may find a subset $\Omega_0 \subseteq \Omega$ with $\mathbf{P}(\Omega_0) = 1$ such that for any $\omega_0 \in \Omega_0$,

$$\begin{cases} V(s, \bar{X}^{t,x;\bar{u}}(s, \omega_0)) = -\bar{Y}^{t,x;\bar{u}}(s, \omega_0), \\ (18), (19), (22), (23), (24), (26) \text{ are satisfied for any} \\ \text{rational } x^1, \text{ and } (\Omega, \mathcal{F}, \mathbf{P}(\cdot | \mathcal{F}_s^t)(\omega_0), W(\cdot) - W(s); \\ u(\cdot)|_{[s,T]}) \in \mathcal{U}^w[s, T]. \end{cases}$$

The first equality of the above is due to the DPP (see Theorem 5.4 of Peng [12]). Let $\omega_0 \in \Omega_0$ be fixed, then for any rational $x^1 \in \mathbf{R}^n$, noting (26), we have

$$\begin{aligned}
&V(s, x^1) - V(s, \bar{X}^{t,x;\bar{u}}(s, \omega_0)) \\
&\leq -Y^{s,x^1;\bar{u}}(s, \omega_0) + \bar{Y}^{t,x;\bar{u}}(s, \omega_0) := -\hat{Y}(s, \omega_0) \\
&= \langle \hat{X}(s, \omega_0), p(s, \omega_0)q^{-1}(s, \omega_0) \rangle + o(|x^1 - \bar{X}^{t,x;\bar{u}}(s, \omega_0)|) \\
&= \langle p(s, \omega_0)q^{-1}(s, \omega_0), X^{s,x^1;\bar{u}}(s) - \bar{X}^{t,x;\bar{u}}(s, \omega_0) \rangle \\
&\quad + o(|x^1 - \bar{X}^{t,x;\bar{u}}(s, \omega_0)|).
\end{aligned} \quad (27)$$

Note that the term $o(|x^1 - \bar{X}^{t,x;\bar{u}}(s, \omega_0)|)$ in the above depends only on the size of $|x^1 - \bar{X}^{t,x;\bar{u}}(s, \omega_0)|$, and it is independent of x^1 . Therefore, by the continuity of $V(s, \cdot)$, we see that (27) holds for all $x^1 \in \mathbf{R}^n$, which by definition (15) proves

$$\{p(s)q^{-1}(s)\} \in D_x^{1,+}V(s, \bar{X}^{t,x;\bar{u}}(s)), \quad \forall s \in [t, T], \quad \mathbf{P}\text{-a.s.}$$

Let us now show $D_x^{1,-}V(s, \bar{X}^{t,x;\bar{u}}(s)) \subset \{p(s)q^{-1}(s)\}$. Fix an $\omega \in \Omega$ such that (27) holds for any $x^1 \in \mathbf{R}^n$. For any $\xi \in D_x^{1,-}V(s, \bar{X}^{t,x;\bar{u}}(s))$, by definition (15) we have

$$\begin{aligned}
0 &\leq \lim_{x^1 \rightarrow \bar{X}^{t,x;\bar{u}}(s)} \left\{ \frac{V(s, x^1) - V(s, \bar{X}^{t,x;\bar{u}}(s))}{|x^1 - \bar{X}^{t,x;\bar{u}}(s)|} \right. \\
&\quad \left. - \frac{\langle \xi, x^1 - \bar{X}^{t,x;\bar{u}}(s) \rangle}{|x^1 - \bar{X}^{t,x;\bar{u}}(s)|} \right\} \\
&\leq \lim_{x^1 \rightarrow \bar{X}^{t,x;\bar{u}}(s)} \frac{\langle p(s)q^{-1}(s) - \xi, x^1 - \bar{X}^{t,x;\bar{u}}(s) \rangle}{|x^1 - \bar{X}^{t,x;\bar{u}}(s)|}.
\end{aligned}$$

Then, it is necessary that

$$\xi = p(s)q^{-1}(s), \quad \forall s \in [t, T], \quad \mathbf{P}\text{-a.s.}$$

Thus, (16) holds. The proof is complete. \square

Remark 3.1 Note that if V is differentiable with respect to x , then (16) reduces to

$$p(s)q^{-1}(s) = V_x(s, \bar{X}^{t,x;\bar{u}}(s)), \quad \forall s \in [t, T], \quad \mathbf{P}\text{-a.s.}, \quad (28)$$

which coincides with the first relation in (3) of Shi [13]. We point out that Theorem 3.1 is a true extension, by which we mean that it is possible to have strict set inclusions in (16). The following example gives such a situation.

Example 3.1 Consider the following controlled SDE ($n = d = 1$):

$$\begin{cases} dX^{t,x;u}(s) = X^{t,x;u}(s)u(s)ds \\ \quad + X^{t,x;u}(s)dW(s), \quad s \in [t, T], \\ X^{t,x;u}(t) = 0, \end{cases} \quad (29)$$

with the control domain being $\mathbf{U} = [0, 1]$. The cost functional is defined as

$$J(t, x; u(\cdot)) := -Y^{t,x;u}(s)|_{s=t}, \quad (t, x) \in [0, T] \times \mathbf{R}^n. \quad (30)$$

with

$$\begin{cases} -dY^{t,x;u}(s) = [X^{t,x;u}(s) - Y^{t,x;u}(s)]ds \\ \quad - Z^{t,x;u}(s)dW(s), \quad s \in [t, T], \\ Y^{t,x;u}(T) = X^{t,x;u}(T). \end{cases} \quad (31)$$

The corresponding generalized HJB equation reads

$$\begin{cases} -v_t(t, x) - \frac{1}{2}x^2v_{xx}(t, x) + x + v(t, x) \\ \quad + \sup_{u \in \mathbf{U}} \{-v_x(t, x)xu\} = 0, \quad (t, x) \in [0, T] \times \mathbf{R}^n, \\ v(T, x) = -x, \quad \forall x \in \mathbf{R}^n, \end{cases} \quad (32)$$

It is not difficult to directly verify that the following function is a viscosity solution to (32):

$$V(t, x) = \begin{cases} -x, & \text{if } x \leq 0, \\ -x(T-t) - x, & \text{if } x > 0, \end{cases} \quad (33)$$

which obviously satisfies (10). Thus, by the uniqueness of the viscosity solution, V coincides with the value function of our problem. Moreover, the adjoint equation writes

$$\begin{cases} -dp(s) = [\bar{u}(s)p(s) - q(s) + k(s)]ds - k(s)dW(s), \\ dq(s) = -q(s)ds, \quad s \in [t, T], \\ p(T) = -q(T), \quad q(t) = 1. \end{cases} \quad (34)$$

Let us consider an admissible control $\bar{u}(\cdot) \equiv 0$ for initial state $x = 0$. The corresponding state under $\bar{u}(\cdot)$ is easily seen to be $\bar{X}^{t,x;\bar{u}}(\cdot) \equiv 0$. By the stochastic verification theorem (see Theorem 9 in [19]), one can check that $(\bar{X}^{t,x;\bar{u}}(\cdot), \bar{u}(\cdot))$ is really optimal. Now let us compare our main result Theorem 3.1 with the one of Shi [13]. In fact, by applying the results of [13], especially (28), we obtain nothing, since $V_x(t, x)$ does not exist along the whole state $\bar{X}^{t,x;\bar{u}}(s), s \in [t, T]$. However, we have

$$\begin{aligned} D_x^{1,-}V(s, \bar{X}^{t,x;\bar{u}}(s)) &= \emptyset, \\ D_x^{1,+}V(s, \bar{X}^{t,x;\bar{u}}(s)) &= [-(T-s) - 1, -1], \end{aligned} \quad (35)$$

and the adjoint process triple is $(p(s), q(s), k(s)) = (-e^{t-s}, e^{t-s}, 0), s \in [t, T]$. Thus the relation (16) holds, which shows that our Theorem 3.1 works.

IV. CONCLUDING REMARKS

In this paper, we have established a nonsmooth version of the connection between the maximum principle and dynamic programming principle, for the stochastic recursive control problem when the control domain is convex. By employing the viscosity solution, the connection is now interpreted as a set inclusion among sub-jet $D_x^{1,-}V(s, \bar{X}^{t,x;\bar{u}}(s))$, super-jet $D_x^{1,+}V(s, \bar{X}^{t,x;\bar{u}}(s))$ and singleton $\{p(s)q^{-1}(s)\}$. This new result has extended the classical one of Shi [13], by eliminating the smoothness assumption on the value function.

This paper is the first part of our recent results on the relationship between maximum principle and dynamic programming principle under the framework of viscosity solutions, for the stochastic recursive optimal control problem. The main result in this paper (Theorem 3.1) is in local form. In the second part, we will deal with its global form, that is, the control domain is not necessarily convex. However, it looks like a difficult problem since the integrability/regularity property of z (the martingale part of the BSDE, which appears in the diffusion coefficient of the forward equation), seems to be not enough in the case when a second-order expression is necessary. In forthcoming research, we will try to overcome this difficulty by using new first- and second-order adjoint equations to deal with the global case.

REFERENCES

- [1] A. Bensoussan, Lectures on stochastic control, *Lecture Notes in Mathematics*, vol. 972, Springer-Verlag, Berlin, 1982.
- [2] J.M. Bismut, An introductory approach to duality in optimal stochastic control. *SIAM Review*, **20(1)**, 62-78, 1978.
- [3] J. Cvitanic, J.F. Zhang, *Contract Theory in Continuous Time Models*, Springer-Verlag, Berlin, 2013.
- [4] D. Duffie, L.G. Epstein, Stochastic differential utility. *Econometrica*, **60(2)**, 353-394, 1992.
- [5] N. El Karoui, S.G. Peng and M.C. Quenez, Backward stochastic differential equations in finance. *Math. Finance*, **7(1)**, 1-71, 1997.
- [6] N. El Karoui, S.G. Peng and M.C. Quenez, A dynamic maximum principle for the optimization of recursive utilities under constraints. *Ann. Appl. Proba.*, **11(3)**, 664-693, 2001.
- [7] W.H. Fleming, R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, 1975.
- [8] N. Nadirashvili, Nonuniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators. *Annali Della Scuola Normale Superiore Di Pisa Classe Di Scienze*, **24(24)**, 537-549, 1997.
- [9] E. Pardoux, S.G. Peng, Adapted solution of a backward stochastic differential equation. *Syst. & Control Lett.*, **14(1)**, 55-61, 1990.
- [10] S.G. Peng, A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation. *Stoch. & Stoch. Reports*, **38(2)**, 119-134, 1992.
- [11] S.G. Peng, Backward stochastic differential equations and applications to the optimal control. *Appl. Math. Optim.*, **27(2)**, 125-144, 1993.
- [12] S.G. Peng, Backward stochastic differential equations—stochastic optimization theory and viscosity solutions of HJB equations. *Topics on Stochastic Analysis*, J. Yan, S. Peng, S. Fang and L. Wu, eds., Beijing, Science Press, 85-138, 1997. (in Chinese)
- [13] J.T. Shi, The relationship between maximum principle and dynamic programming principle for stochastic recursive optimal control problems and applications to finance. *Proc. 29th Chinese Control Conf.*, 1535-1540, July 29-31, Beijing, China, 2010.
- [14] J.T. Shi, Z.Y. Yu, Relationship between maximum principle and dynamic programming for stochastic recursive optimal control problems and applications. *Math. Prob. Engin.*, Vol. 2013, Article ID 285241, 12 pages.
- [15] G.C. Wang, Z. Wu, The maximum principle for stochastic recursive optimal control problems under partial information. *IEEE Trans. Autom. Control*, **54(6)**, 1230-1242, 2009.
- [16] Z. Wu, A general maximum principle for optimal control problems of forward-backward stochastic control systems. *Automatica*, **49(5)**, 1473-1480, 2013.
- [17] W.S. Xu, Stochastic maximum principle for optimal control problem of forward and backward system. *J. Aust. Math. Soc., Ser. B*, **37(2)**, 172-185, 1995.
- [18] J.M. Yong, X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.
- [19] L.Q. Zhang, Stochastic verification theorem of forward-backward controlled systems for viscosity solutions. *Syst. & Control Lett.*, **61(5)**, 649-654, 2012.
- [20] X.Y. Zhou, Maximum principle, dynamic programming, and their connection in deterministic control. *J. Optim. Theory Appl.*, **65(2)**, 363-373, 1990.